Eye tracking with the adaptive optics scanning laser ophthalmoscope

CONFERENCE PAPER · MARCH 2010

DOI: 10.1145/1743666.1743714 · Source: DBLP

3 AUTHORS, INCLUDING:

Scott B Stevenson

University of Houston

Girish Kumar

University of California, Berkeley

Eye Tracking with the Adaptive Optics Scanning Laser Ophthalmoscope

Scott B. Stevenson[1], Austin Roorda[2] and Girish Kumar[1]
1 University of Houston College of Optometry, 2 University of California-Berkeley School of Optometry

Abstract

Recent advances in high magnification retinal imaging have allowed for visualization of individual retinal photoreceptors, but these systems also suffer from distortions due to fixational eye motion. Algorithms developed to remove these distortions have the added benefit of providing arc second level resolution of the eye movements that produce them. The system also allows for visualization of targets on the retina, allowing for absolute retinal position measures to the level of individual cones. This paper will describe the process used to remove the eye movement artifacts and present analysis of their spectral characteristics. We find a roughly 1/f amplitude spectrum similar to that reported by Findlay (1971) with no evidence for a distinct tremor component.

Keywords: Eye tracking technologies, retinal imaging, ocular tremor.

1 Introduction / Overview

A wide array of technologies exist for tracking movements of the eye, and each has particular useful features that others may lack. In particular, tracking methods vary considerably in the range of eye movement amplitudes they can usefully record, from the smallest movements of steady fixation to the full range of gaze (Figure 1). This paper describes a method for tracking small eye movements using a very high magnification retinal imaging system that allows visualization of individual retinal cone photoreceptors. The device is an Adaptive Optics Scanning Laser Ophthalmoscope (AOSLO), which forms images of the retina by raster scan of a laser spot at a line rate of 16 kHz and a frame rate of typically 30 Hz. The high magnification is achieved by measuring the eye's natural optical aberrations with a wavefront sensor, and then correcting those aberrations with a deformable mirror. In this way the eye becomes diffraction limited over a relatively large (5-6 mm) pupil and features of the retina on the scale of a few microns can be imaged [Roorda et al. 2002]. The AOLSO has the added benefit that stimuli can be delivered to the retina with the same high spatial resolution, with direct visualization of the retinal image location of the stimuli [Arathorn et al. 2007: Raghunandan et al. 2008]

Eye motion in this system produces dramatic distortions of the scanned images, however. While the horizontal scan rate of 16 kHz is rapid relative to most eye motion, the vertical scan is a relatively slow rate of 30 Hz. Saccades and drifts of the eye cause the raster to jitter about on the retina and the resulting recorded images show compression, expansion, and shearing artifacts similar to those commonly seen in photocopies when the original is moved during a scan (Figure 2). In raw images of the retina, such distortions are usually less obvious on inspection, but can sometimes produce dramatic artifacts in the image. Removal of these artifacts by offline processing of recorded AOSLO video has the side benefit of revealing the eye motion that produce them, and so the AOSLO is both a retinal imager and an eye tracker.

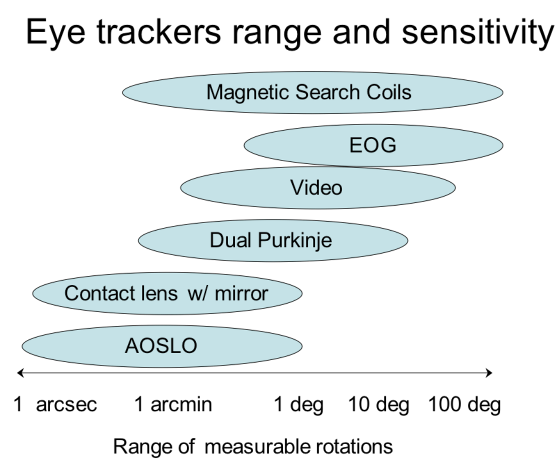

Figure 1. Sensitivity ranges for methods of eye tracking.

Recovery of and correction for eye motion from Scanning Laser Ophthalmoscope images has been reported previously by a number of research groups [Lakshminarayanan et al. 1992: Mirisita and Yagi 2001: Mulligan 1997: Ott and Daunicht 1992: Ott and Eckmiller 1989: Schuchard and Raasch 1992: Stetter et al. 1996]. Scanning laser ophthalmoscopes are built in to some commercial instruments for use as eye position monitors, such as the Heidelberg Spectralis Optical Coherence Tomography (OCT) scanner (http://www.heidelbergengineering.com). Tracking with wide field SLO images typically involves identification of a particular feature, such as a blood vessel bifurcation, and then following that feature from frame to frame. The Adaptive Optics SLO has the advantage that the entire field of view has high contrast trackable image content, so that eye motion can be tracked within a frame at high temporal rates. The high magnification of the system allows tracking of microsaccades and other fixation eye movement components to a resolution of about an arc second.

2 Image Registration

Recovery of eye motion from an AOSLO movie requires registration of image content within each frame to a standard refer-

ETRA 2010, Austin, TX, March 22 – 24, 2010.
© 2010 ACM 978-1-60558-994-7/10/0003 $10.00

ence image. The reference image is assumed to reflect an undistorted retinal image, and so shifts of image content at any point in an individual frame reflect the motion of the eye at the point in time when the retina was being scanned. Undistorted images of the retina are not typically available from a separate source, so they must be built up from all the frames in a movie under the assumption that the average motion of the eye is zero over the several seconds of data collected in a typical movie. Thus, the average distortion is zero.

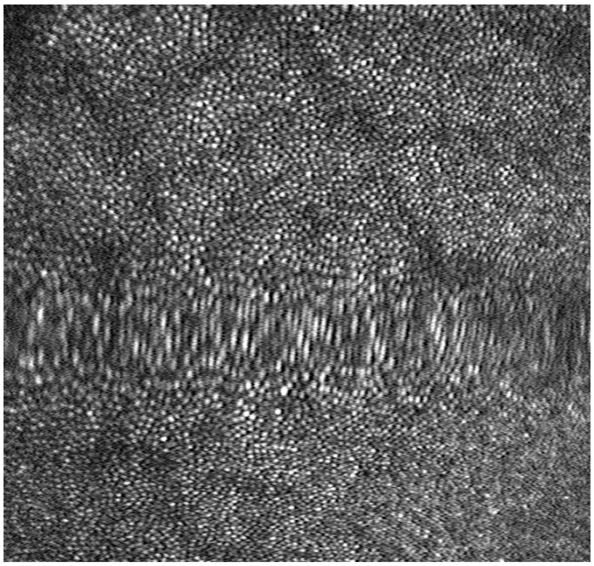

Figure 2. Single frame from an AOSLO scan of a human retina The distortion in the center of the frame is due to a downward vertical saccade.

In practice, one can produce a much better reference frame by removing frames with obvious distortion as in Figure 2 or frames which correlate poorly to others due to blinks or large saccades. Our approach is to build a reference frame iteratively, starting with an average made from a few good frames, then correlating each frame to that reference to determine where saccades have occurred. Frames without saccades are then combined to produce a more refined reference image, which is in turn used to register individual horizontal strips of each frame to the reference. The final time resolution of the extracted eye trace is determined by the number of strips used per frame multiplied by the frame rate of 30 Hz. In principle, each scan line can be registered to the reference frame to yield a temporal resolution of 16 kHz, but in practice we typically use wider strips and analyze at 500 to 1000 Hz.

If the reference frame has any residual distortion, in general this will show up as a 30 Hz (and harmonics thereof) artifact in the eye movement trace. Such frame rate artifacts may also show up if there is an overall slow torsion change through the recording period or if the scanner itself changes amplitude slightly. While these artifacts do occur, in practice they are small and can be removed by calculating and subtracting the average motion at the frame rate.

Registration of image content to a reference image involves computing a two dimensional cross-correlation, which we implement in Matlab code using the function FFT2. The quality of a match is assessed by comparing the peak correlation value to the next highest peak in the correlogram, providing a confidence measure that a unique match has been found. The search range in the final registration step is constrained by use of a cubic spline fit to the motion trace extracted in prior, more coarse matching steps. Subpixel resolution is achieved by fitting a 2D cubic spline function to the correlation peak in the final stage.

While we use two dimensional correlation to extract horizontal and vertical motion, it is also possible to extract torsional motion by comparing the vertical shift on the left and right halves of each strip of video. Previous studies have shown that torsional movements may be as much as twice the amplitude of horizontal and vertical movements during fixation [van Rijn et al. 1994]. A shift of one pixel difference corresponds to about 10 arc minutes of torsion, so the AOSLO is much less sensitive to torsion compared to horizontal and vertical movements and generally we do not extract and correct this component.

3 Validation of the eye motion extraction

Once the motion trace has been extracted, a stabilized video is synthesized by creating frames in which each raster line is placed at the registered location, resulting in undistorted retinal structure in the movie. The principle validation of the accuracy of the eye motion extraction comes from the stabilized movie itself. If individual cone photoreceptors are stationary in the video, it indicates that the eye motion artifacts have been successfully measured and removed. For most videos, the residual motion is considerably less than the one arc minute diameter of a cone.

We also have directly compared the motion extracted from AOSLO video to simultaneous recordings made with a dual Purkinje image eye tracker. [Cornsweet and Crane 1973] Figure 3 shows superimposed vertical eye position traces from a 4 sec-

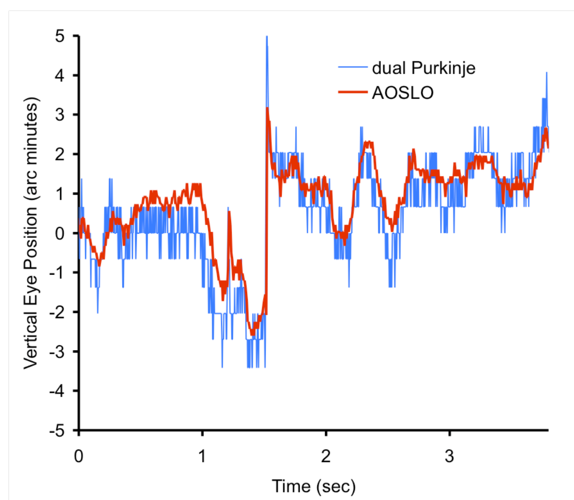

Figure 3. Comparison of eye movement traces collected simultaneously from a dual Purkinje image eye tracker (blue) and recovered from movies made with the AOSLO retinal imager (red). The dPi tracker signals were digitized at a resolution of one arc minute, its nominal noise level. The AOSLO traces were recovered at one pixel resolution, in this case 7 arc seconds. Subpixel recovery from AOSLO videos allows resolution of one arc second or better.

ond segment of recording. The RMS difference between the two methods was roughly one arc minute, which is the expected noise level for the dual Purkinje tracker. Note that the lens wobble artifact that is commonly seen in dual Purkinje records [Deubel and Bridgeman 1995] is present also in the AOSLO record to a lesser degree. The motion of the eye's crystalline lens around the time of a saccade produces a shift of the retinal image as predicted from ray tracing eye models.

For comparison to the motion of a real eye, we also imaged an artificial, stationary eye to provide an estimate of the noise level in image acquisition and motion extraction. The artificial eye consisted of a plus lens and a piece of paper, mounted rigidly to the AOSLO platform. Figure 4 shows eye motion traces recovered from a human eye in steady fixation, and traces from the model eye. The model eye traces are steady to within one pixel of the video, here about 15 arc seconds.

A final step in validating the method for extracting eye motion was to generate synthetic movies with known imposed motion and then compare the recovered motion to the original, imposed motion. A single image of retinal structure was used as the basis for the movie, and along with the distortions, we added luminance noise to each frame to determine the robustness of the extraction to this type of noise. Even when the added noise was a substantial fraction of the full 0 to 255 luminance range of the image, the mean absolute difference was just a few seconds of arc. For most of our recorded videos, the standard deviation of luminance is around 20, where the algorithm performed at the 2 arc second level. Thus, our extraction method is robust to added luminance noise.

4. Analysis of fixation eye movements.

The high sensitivity of the AOSLO eye tracking system makes it ideal for studying miniature eye movements, and in particular for examining the drift and tremor components described previously. Adler and Fliegelman [1934] described fixation movements as having three distinct components: microsaccades, drifts and tremor, with the last being high frequency and less than one arc minute in amplitude. Many reports describe a distinct tremor component with a dominant frequency at between 50 and 100 Hz. [Ratliff and Riggs 1950; Ditchburn and Ginsborg 1953; Riggs et al. 1954; Matin 1964; Steinman et al 1982; Bolger et al 1999], but at least two reports that analyzed the spectrum of fixation eye movements describe an overall reciprocal relation

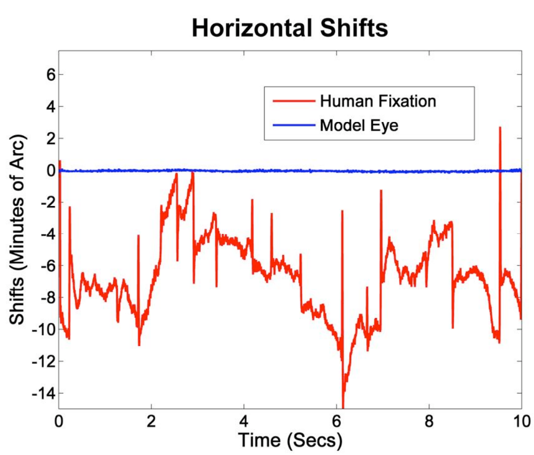

Figure 4. Horizontal eye position traces from a human eye and a model eye, extracted from AOSLO video.

between amplitude and frequency with an inflection [Findlay 1971] or relative increase in amplitude [Eizenman et al 1999] at around 100 Hz.

For comparison to previous studies of miniature eye movements, and to clarify and extend these observations of the tremor component, we performed at 1400 Hz eye movement extraction from 200 msec long segments of saccade-free steady fixation, recorded from four subjects. We also extracted motion from video

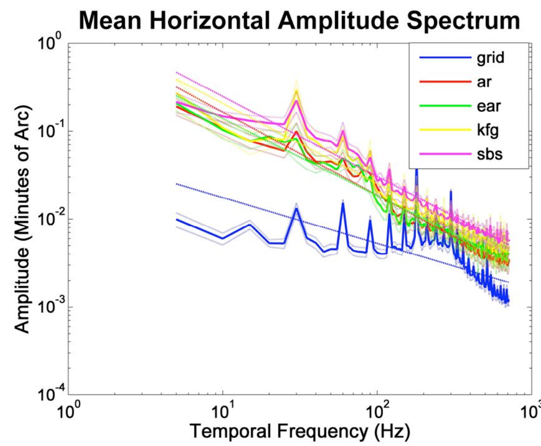

Figure 5. Amplitude spectrum calculated with Matlab's FFT function using eye traces extracted from AOSLO video. Segments without saccades were chosen from several videos of subjects fixating a 16 arcmin square target. Each analyzed segment lasted 200 msec and eye motion was extracted at 1400 Hz, producing a spectrum from 5 to 700 Hz. The bottom trace labeled 'grid" was extracted from a stationary, model eye and had an amplitude of around 0.01 arc minutes (0.6 arc seconds) over most of the frequency range tested. All four subjects showed generally a 1/f relationship between amplitude and frequency over the entire spectrum (fitted lines), with a small relative increase between 50 and 100 Hz. Note that peaks at 30 Hz and higher harmonics are an artifact and show up in the artificial eye as well as the human eyes.

of a grid pattern, imaged in place of the eye, in order to estimate the noise level across frequency.

Amplitude spectra obtained using the Matlab FFT function on these segments are plotted in Figure 5 for four subjects, who show good agreement. The spectrum obtained from the model eye ("grid") is shown for comparison. Noise level from the model eye is less that one arc second across most of the spectrum, while all four subjects show a reciprocal relationship between amplitude and frequency that is well fit by a slope of -1 on log log coordinates. A small increase relative to the 1/F relationship is apparent between 50 and 100 Hz, the usual frequencies associated with tremor. Narrow peaks at harmonics of the 30 Hz frame rate are evident in both the human traces and the model eye traces, indicating that they are artifacts of the scanning.

Examination of these spectra confirms the 1/F finding by Findlay [1971] for frequencies up to 200 Hz. However, our analysis extends the 1/F relation to 700 Hz., in contrast to Findlay who found a steeper slope between 100 and 200 Hz. One difference between the studies is that his analysis required use of a model to separate saccadic and drift/tremor components, whereas we selected segments that were saccade free.

While there does seem to be a relative increase in amplitude in the 50 to 100 Hz range, as reported by Eizenman et al [1999] and others, it is not obvious that a clear "tremor" component exists at those frequencies. Perhaps the most accurate characterization of fixation eye movements would be that it is a scale-invariant random walk pattern punctuated occasionally by distinct microsaccadic jerks.

Supported by NSF AST-9876783 through the UC Santa Cruz Center for Adaptive Optics

References

ADLER FH & FLIEGELMAN 1934. Influence of fixation on the visual acuity. *Arch. Ophthalmology 12*, 475.

ARATHORN D. W., YANG Q., VOGEL C. R., ZHANG Y., TIRUVEEDHULA P., AND ROORDA A., 2007. Retinally stabilized cone-targeted stimulus delivery, *Opt. Express* 15, 13731-13744

BOLGER C, BOJANIC S, SHEAHAN NF, COAKLEY D, MALONE JF 1999. Dominant frequency content of ocular microtremor from normal subjects *Vision Research* 39, 1911-1915.

CORNSWEET TN, CRANE HD. 1973. Accurate two-dimensional eye tracker using first and fourth Purkinje images. *J Opt Soc Am. 63*, 921-8.

CORNSWEET TN. 1958. New technique for the measurement of small eye movements. *J Opt Soc Am 48*, 808-811.

DEUBEL, H. AND BRIDGEMAN, B. 1995. Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research 35*, 529-538.

DITCHBURN, R. W., & GINSBORG, B. L. (1953). Involuntary eye movements during fixation. *Journal of Physiology*, 119, 1 – 7.

EIZENMAN M, HALLETT PE, FRECKER RC. 1985. Power spectra for ocular drift and tremor. *Vision Research 25*, 1635-40

EIZENMAN, M., HALLETT, P. E, & FRECKER, R. C. (1985). Power spectra for ocular drift and tremor. *Vision Research*, 25(11), 1635 – 1640.

FINDLAY JM (1971) Frequency analysis of human involuntary eye movement. *Kybernetic* 8(6):207-14.

LAKSHMINARAYANAN V, KNOWLES RA, ENOCH JM, & VASUVEDAN R 1992. Measurement of fixational stability while performing a hyperacuity task using the scanning laser ophthalmoscope: preliminary studies. *Clinical Vision Sciences 7*, 557-563.

MATIN, L. (1964). Measurement of eye movements by contact lens techniques. *Journal of the Optical Society of America*, 54, 1008.

MORISITA M, YAGI T. 2001. The stability of human eye orientation during visual fixation and imagined fixation in three dimensions. *Auris Nasus Larynx. 283*, 301-304.

MULLIGAN, JB, 1997. Recovery of Motion Parameters from Distortions in Scanned Images. *Proceedings of the NASA Image Registration Workshop IRW97.*, NASA Goddard Space Flight Center, MD

OTT D & DAUNICHT WJ 1992. Eye movement measurement with the scanning laser ophthalmoscope. *Clinical Vision Sciences. 7*, 551-556.

OTT D & ECKMILLER R 1989. Ocular torsion measured by TV- and scanning laser ophthalmoscopy during horizontal pursuit in humans and monkeys. *Investigative Ophthalmology and Visual Sciences 30*, 2512-2520.

RAGHUNANDAN, A., FRASIER, J., POONJA, S., ROORDA, A., & STEVENSON, S. B. 2008. Psychophysical measurements of referenced and unreferenced motion processing using high-resolution retinal imaging. *Journal of Vision*, 8 (14), 1-11.

RATLIFF, F., & RIGGS, L. A. (1950). Involuntary motions of the eyes during monocular fixation. *Journal of Experimental Psychology*, 40, 687.

RIGGS LA, ARMINGTON JC & RATLIFF F. 1954. Motions of the retinal image during fixation. *J Opt Soc Am 44*, 315-321.

RIGGS, L. A, ARMINGTON, J. C., & RATCLIFF, F. (1954). Motions of the retinal image during fixation. *Journal of the Optical Society of America,* 44(4), 315 – 321.

RIGGS, L. A. & NIEHL, E. W. 1960. Eye movements recorded during convergence and divergence. *J Opt Soc Am 50*, 913-920.

ROORDA, A., ROMERO-BORJA, F., DONNELLY III, W.J., QUEENER, H., HEBERT, T.J., CAMPBELL, M.C.W. 2002. Adaptive Optics Scanning Laser Ophthalmoscopy. *Opt. Express 10*, 405-412.

SCHUCHARD RA & RAASCH TW 1992. Retinal locus for fixation: pericentral fixation targets. *Clinical Vision Sciences. 7*, 511-520.

STEINMAN R. M., CUSHMAN, W. B., & MARTINS, A. J. (1982). The precision of gaze. *Human Neurological Biology*, 1, 97 – 109

STEINMAN RM, HADDAD GM, SKAVENSKI AA, WYMAN D. 1973. Miniature eye movement. *Science 1818*, 10-9.

STETTER M, SENDTNER RA, TIMBERLAKE GT. 1996 A novel method for measuring saccade profiles using the scanning laser ophthalmoscope. *Vision Research 36*, 1987-94

STEVENSON SB, & ROORDA A., 2005 "Correcting for miniature eye movements in high resolution scanning laser ophthalmoscopy" *in Ophthalmic Technologies XV*, edited by Fabrice Manns, Per Söderberg, Arthur Ho, Proceedings of SPIE Vol. 5688A (SPIE, Bellingham, WA, 2005).

VAN RIJN LJ, VAN DER STEEN J, & COLLEWIJN H 1994. Instability of ocular torsion during fixation: cyclovergence is more stable than cycloversion. *Vision Research 34*, 1077-1087